

# Generalized Lineage-Aware Temporal Windows: Supporting Outer and Anti Joins in Temporal-Probabilistic Databases

Katerina Papaioannou\*, Martin Theobald†, Michael Böhlen\*

\* *Department of Computer Science – University of Zurich*  
{papaioannou, boehlen}@ifi.uzh.ch

† *Faculty of Science, Technology & Communication – University of Luxembourg*  
martin.theobald@uni.lu

**Abstract**—The result of a temporal-probabilistic (TP) join with negation includes, at each time point, the probability with which a tuple of a positive relation  $p$  matches *none* of the tuples in a negative relation  $n$ , for a given join condition  $\theta$ . TP outer and anti joins thus resemble the characteristics of relational outer and anti joins also in the case when there exist time points at which input tuples from  $p$  have non-zero probabilities to be *true* and input tuples from  $n$  have non-zero probabilities to be *false*, respectively. For the computation of TP joins with negation, we introduce *generalized lineage-aware temporal windows*, a mechanism that binds an output interval to the lineages of all the matching valid tuples of each input relation. We group the windows of two TP relations into three disjoint sets based on the way attributes, lineage expressions and intervals are produced. We compute all windows in an incremental manner, and we show that pipelined computations allow for the direct integration of our approach into PostgreSQL. We thereby alleviate the prevalent redundancies in the interval computations of existing approaches, which is proven by an extensive experimental evaluation with real-world datasets.

## I. INTRODUCTION

Join operations with negation are performed for a positive relation  $p$ , a negative relation  $n$  and a  $\theta$  condition that determines the tuples that match. In conventional databases, joins with negation disqualify an input tuple of the positive relation if its attributes match the attributes in a tuple of the negative relation. In temporal databases, the existence of a matching tuple in  $n$  does not disqualify the tuple of  $p$  itself but timepoints at which it is valid [1], [2]. In probabilistic databases, where tuples have a probability to be true or false, the existence of a matching tuple in  $n$  only reduces the probability with which a tuple is included in the output [3], [4].

The result of a temporal-probabilistic join with negation includes, at each time point, the probability with which a tuple of the positive relation  $p$  matches no tuple in the negative relation  $n$  for a predicate  $\theta$ . Firstly, it includes output tuples that span subintervals when only tuples of  $p$  are valid. In such cases, output intervals might be determined by starting or ending points of input tuples that are not valid during the output interval. Secondly, TP joins with negation produce outputs that indicate, at each time point, the probability of a

tuple  $\tilde{p}$  in  $p$  not matching any valid tuple in  $n$  because all of them are false. In this case, an output interval  $T$  is determined based on the starting and ending points of  $\tilde{p}$  and of the tuples of  $n$  that are valid over  $T$  and match  $\tilde{p}$  for  $\theta$ .

a (wantsToVisit)					b (hotelAvailability)				
Name	Loc	$\lambda$	$T$	$p$	Hotel	Loc	$\lambda$	$T$	$p$
Ann	ZAK	$a_1$	[2,8]	0.7	hotel <sub>3</sub>	SOR	$b_1$	[1,4]	0.9
Jim	WEN	$a_2$	[7,10]	0.8	hotel <sub>2</sub>	ZAK	$b_2$	[5,8]	0.6
					hotel <sub>1</sub>	ZAK	$b_3$	[4,6]	0.7

(a) Temporal-probabilistic base relations

$Q = a \bowtie_{\theta}^{\text{tp}} b, \theta : a.\text{Loc} = b.\text{Loc}$					
Name	Loc	Hotel	$\lambda$	$T$	$p$
Ann	ZAK	-	$a_1$	[2,4]	0.70
Ann	ZAK	hotel <sub>1</sub>	$a_1 \wedge b_3$	[4,6]	0.49
Ann	ZAK	hotel <sub>2</sub>	$a_1 \wedge b_2$	[5,8]	0.42
Ann	ZAK	-	$a_1 \wedge \neg b_3$	[4,5]	0.21
Ann	ZAK	-	$a_1 \wedge \neg(b_3 \vee b_2)$	[5,6]	0.084
Ann	ZAK	-	$a_1 \wedge \neg b_2$	[6,8]	0.28
Jim	WEN	-	$a_2$	[7,10]	0.80

(b) Temporal-probabilistic tuple-based query

**Fig. 1:** Temporal-probabilistic database example

*Example 1:* Consider a booking website (Figure 1) that archives prediction data over time. Table **a** records data related to the locations that the clients want to visit, according to their searches. Table **b** records data regarding the availability of the hotels registered in the website, considering the busy periods in each location and the rate at which each hotel gets booked. This archive corresponds to a temporal-probabilistic database. Tuple ('Jim, WEN',  $a_2$ , [7,10), 0.8) captures that, at each day from the 7<sup>th</sup> to the 10<sup>th</sup> of the month, 'Jim wants to visit Wengen' with probability 0.8. The website makes a prediction for each time point and there is no other tuple in **a** that predicts the probability of 'Jim visiting Wengen' over an interval overlapping with [7,10). In order to manage supply and demand, we determine the probability with which the client will find available accommodation at their preferred location, at each time point. The corresponding query is  $Q$

$= \mathbf{a} \bowtie_{\theta}^{\text{TP}} \mathbf{b}$  ( $\theta : \mathbf{a}.\text{Loc} = \mathbf{b}.\text{Loc}$ ), i.e., a temporal-probabilistic outer join with equality on the locations.

The answer tuple ('Ann, ZAK, hotel<sub>1</sub>',  $a_1 \wedge b_3$ , [4,6), 0.49) expresses that, with probability 0.49, Ann wants to visit Zakynthos ( $a_1$ ) and stay at hotel<sub>1</sub> in Zakynthos ( $b_3$ ) during interval [4,6). It is valid over the intersection of the intervals of tuples  $a_1$  and  $b_3$  and it is true when both these tuples are true. Answer tuple ('Ann, ZAK, -',  $a_1$ , [2,4), 0.7) expresses that, with probability 0.7, Ann wants to visit Zakynthos ( $a_1$ ) but there is no hotel available to stay there. Although the lineage and the output probability are both determined by tuple  $a_1$ , i.e., the only tuple valid during [2,4), the interval of this output tuple is influenced by the starting point of tuple  $b_3$ , a tuple not valid over [2,4). Over the interval [5,6) there is 0.084 probability that Ann wants to visit Zakynthos but finds no accommodation. According to answer tuple ('Ann, ZAK, -',  $a_1 \wedge \neg(b_3 \vee b_2)$ , [5,6), 0.084), during [5,6), the output is influenced by more than a pair of input tuples. Although all tuples are valid over [5,6), this tuple is true when 'Ann visits Zurich' ( $a_1$  is true) but also when neither hotel<sub>1</sub> nor hotel<sub>2</sub> are available during [5,6) ( $b_3$  and  $b_2$  are false).

TP set-difference is the only temporal-probabilistic operation with negation that has been investigated [5]. Since set-operations combine only tuples with equal non-temporal attributes, simplified structures can be used. Specifically, only one tuple of each relation is valid at each time point, which allows for solutions with linearithmic complexity. For TP outer joins and TP anti join, multiple tuples of the negative relation might be valid over an output interval and input tuples with non-temporal attributes that are not pairwise equal might be combined to form an output tuple. Moreover, TP outer joins combine the characteristics of TP joins with and without negation: at each time point, two outcomes are possible since the same tuples can be *true* or *false*.

	$F_r$	$F_s$	$\lambda_r$	$\lambda_s$	$T$
$w_1$	'Ann, ZAK'	-	$a_1$	-	[2,4)
$w_2$	'Jim, WEN'	-	$a_2$	-	[7,10)

(a) Unmatched Windows

	$F_r$	$F_s$	$\lambda_r$	$\lambda_s$	$T$
$w_3$	'Ann, ZAK'	'hotel <sub>1</sub> , ZAK'	$a_1$	$b_3$	[4,6)
$w_4$	'Ann, ZAK'	'hotel <sub>2</sub> , ZAK'	$a_1$	$b_2$	[5,8)

(b) Overlapping Windows

	$F_r$	$F_s$	$\lambda_r$	$\lambda_s$	$T$
$w_5$	'Ann, ZAK'	-	$a_1$	$b_3$	[4,5)
$w_6$	'Ann, ZAK'	-	$a_1$	$b_3 \vee b_2$	[5,6)
$w_7$	'Ann, ZAK'	-	$a_1$	$b_2$	[6,8)

(c) Negating Windows

**Fig. 2:** Generalized lineage-aware temporal windows of relations **a** and **b** (Fig. 1a) for the  $\theta$ -condition  $\mathbf{a}.\text{Loc}=\mathbf{b}.\text{Loc}$

## Outline & Contributions.

- We introduce *generalized lineage-aware temporal windows* to produce output tuples for input pairs with different

non-temporal attributes and for cases when multiple input tuples are valid. Given a  $\theta$ -condition and two TP relations, we group windows into three disjoint sets: the *unmatched*, the *overlapping* and the *negating windows*. An output tuple is formed for each window using the appropriate lineage-concatenation functions and we express the result of TP joins with negation using the three sets.

- We introduce the algorithms  $\text{LAWA}_U$  and  $\text{LAWA}_N$  for the computation of unmatched and negating windows, respectively. Recording the lineages of the tuples valid in each input relation over an output interval and keeping them decoupled until the formation of output tuples, allows for the computation of unmatched and negating windows based on the overlapping ones. Thus, redundant interval comparisons due to the repetition of basic steps are avoided and the runtime required for the computation of outer joins and anti join improves by two orders of magnitude.
- We conduct extensive experiments using real datasets to compare our approach for the computation of TP outer joins and TP anti join with existing state of the art approaches. Our approach is integrated in PostgreSQL and exhibits a lower runtime while being scalable.

The remainder of this paper is organized as follows. Section II provides an overview of related works on temporal and probabilistic databases with a focus on outer joins and anti join. Section III discusses the TP data model and its query semantics. Section IV discusses the impact of negation in TP joins. Section V introduces generalized lineage-aware temporal windows and groups them into three disjoint sets. Section VI introduces two algorithms for the computation of the different window sets while section VII presents a comprehensive performance study that compares our implementation with existing approaches. Section VIII concludes the paper.

## II. RELATED WORK

We review related approaches from temporal and probabilistic databases and explain their limitations in terms of supporting TP outer joins and anti join.

**Temporal-Probabilistic Operations.** Dylla et al. [6] introduced a closed and complete TP database model, coined TPDB, based on existing temporal and probabilistic models. Query processing is performed in two steps. The first step, grounding, evaluates a chosen deduction rule (formulated in Datalog with additional time variables and temporal predicates) and computes the lineage expressions of the deduced tuples. The second step, deduplication, removes the duplicates that could occur in the grounding step by adjusting intervals. The grounding step performs pairwise tuple-comparisons. Subintervals that are present in only one of the two input relations, i.e., during which no tuple of the other relation is valid, cannot be produced.

**TP Operations with negation.** Set-difference is the only TP operation with negation that has been investigated [5]. For its computation, Papaioannou et al. introduced *lineage-aware*

*temporal windows*, a mechanism that binds an output interval with the lineage of the tuple in each input relation that includes fact  $F$  and that is valid during the interval. *Lineage-aware temporal windows* eliminate redundant interval comparisons and additional joins for the formation of lineage expressions in TP set operations. The starting and ending points of the interval that the window spans are computed via a comparison of the starting and ending points of input tuples that are valid but also of neighboring tuples. Thus, they are useful for output intervals that are not equal to the overlap of a pair of valid tuples. However, they are tailored to cases when one tuple of each input relation is valid and when the input tuples have the same non-temporal attributes. In TP joins with negation, input tuples with different non-temporal attributes are combined and multiple tuples of an input relation can be valid over an interval and need to be included in the lineage of an output tuple.

**Temporal Joins.** In temporal databases, the result of a temporal outer join  $op^T$  is defined as the result of applying  $op$  over a sequence of atemporal instances (the so-called snapshots) of the input relations—a key concept in temporal databases termed *snapshot reducibility* [7], [8], [9]. Maximal intervals are produced by merging consecutive time points to which the same input tuples have contributed (*change preservation*). Dignös et al. [10], [11] use *data lineage* to guarantee change preservation for all relational operations under a sequenced semantics. For the computation of joins, they introduce the *alignment* operator. The alignment  $\Phi(r, s)$  of a relation  $r$  based on another relation  $s$  replicates the tuples of  $r$  and assigns new time intervals to them. The new intervals are obtained by splitting the original intervals of  $r$  based on tuples of  $s$  with which they overlap. The valid tuples of both relations that contribute to an adjusted interval are not recorded. This is the reason why the alignment of both relations is required as well as the application of  $op$  to produce all output tuples [10], [11]. Using this approach in a TP context, other than the overhead and redundancy of aligning both relations, the input tuples must also be adjusted in groups and not only in pairs for the cases when valid tuples are *false*. Combining adjustment both in pairs and in groups multiple times in the same query incurs redundant comparisons and recomputation of intermediate results.

*Sweeping-based approaches* have been widely used for the computation of overlap joins [12], [13] in temporal settings. A sweepline moves over all start and end points of tuples, and determines, for each time point, the tuples of both input relations that are valid. These approaches are tailored to compute efficiently the overlap join but are not suitable for the computation of the class of operations discussed in this paper. First, the overlapping intervals computed in these approaches only correspond to a part of the result of a TP outer join while they are not included in the result of a TP anti join. Second, they generally do not consider join conditions on the non-temporal attributes limiting the types of queries they could be used for.

**Probabilistic Joins.** In probabilistic databases, the result of a

probabilistic operation  $op^p$  is defined as the result of applying  $op$  over the set of all possible instances of the input relations. The Trio system [14] was among the first to recognize *data lineage*, in the form of a Boolean formula, as a means to capture the possible instances at which an output tuple is valid. In an effort to provide a *closed and complete* representation model for uncertain relational data, they introduced *Uncertainty and Lineage Databases* (ULDBs) [15]. The algebraic operators are modified to compute the lineage of the result tuples in a ULDB, thus capturing all information needed for computing query answers and their probabilities. Fink et al. [16], [17] reduced the computation of probabilistic algebraic operations to conventional operations so that these can be performed using a DBMS, rather than by an application layer built on top of it. In all these works, the focus is restricted to select-project join queries. Probabilistic anti join, expressed with the NOT EXISTS predicate in SQL, has been explored by Wang et al. [4]. It has been integrated in MystiQ by breaking the initial query into positive and negative subqueries that are separately evaluated and then combined. Incorporating interval computation with predicates in these approaches is possible but does not comply with all the requirements of TP operations with negation.

### III. BACKGROUND

We denote a **temporal-probabilistic schema** by  $R^{Tp}(F, \lambda, T, p)$ , where  $F = (A_1, A_2, \dots, A_m)$  is an ordered set of attributes, and each attribute  $A_i$  is assigned to a fixed domain  $\Omega_i$ .  $\lambda$  is a Boolean formula corresponding to a lineage expression.  $T$  is a *temporal attribute* with domain  $\Omega^T \times \Omega^T$ , where  $\Omega^T$  is a finite and ordered set of *time points*.  $p$  is a *probabilistic attribute* with domain  $\Omega^p = (0, 1] \subset \mathbb{R}$ . A **temporal-probabilistic relation**  $r$  over  $R^{Tp}$  is a finite set of tuples. Each tuple  $r \in r$  is an ordered set of values from the appropriate domains. The value of attribute  $A_i$  of  $r$  is denoted by  $r.A_i$ . The conventional attributes  $F = (A_1, A_2, \dots, A_m)$  of tuple  $r$  form a *fact*, and we write  $r.F$  to denote the fact  $f$  captured by tuple  $r$ . For example, base tuple ('Ann, ZAK',  $a_1$ , [2, 8), 0.7) of relation  $a$  (see Fig. 1a) includes the fact  $a_1.F = ('Ann, ZAK')$ , the lineage expression  $a_1.\lambda = a_1$ , the time interval  $a_1.T = [2, 8)$ , and the probability value  $a_1.p = 0.7$ . The temporal-probabilistic annotations of the schema express that (i)  $a_1 = true$  with probability  $a_1.p$  for every time point in  $a_1.T$ , (ii)  $a_1 = false$  with probability  $1 - a_1.p$  for every time point in  $a_1.T$ , (iii) and  $a_1$  is always *false* outside  $a_1.T$ . By following conventions from [6], [11], [10], [18], we assume duplicate-free input and output relations. Formally, a temporal-probabilistic relation  $r$  is **duplicate-free** iff  $\forall r, r' \in r (r \neq r' \Rightarrow r.F \neq r'.F \vee r.T \cap r'.T = \emptyset)$ . In other words, the intervals of any two tuples of  $r$  with the same fact  $f$  do not overlap.

A **lineage expression**  $\lambda$  is a Boolean formula, consisting of tuple identifiers and the three Boolean connectives  $\neg$  ("not"),  $\wedge$  ("and") and  $\vee$  ("or"). Tuple identifiers represent Boolean random variables among which we assume independence [6], [18], [19]. For a base tuple  $r$ ,  $r.\lambda$  is an atomic expression

consisting of just  $r$  itself. For a result tuple  $\tilde{r}$  derived from one or more TP operations,  $\tilde{r}.\lambda$  is a Boolean expression as defined above. The probability of a result tuple is computed via a probabilistic valuation of the tuple's lineage expression, using either exact (see, e.g., [19], [20], [21]) or approximate (see, e.g., [22], [23], [24], [25], [26]) algorithms. For example, in the result relation of Fig. 1b, the lineage  $a_1 \wedge \neg b_3$  yields a marginal probability of  $0.7 \cdot (1 - 0.7) = 0.21$  by assuming independence among the base tuples  $a_1$  and  $b_3$  (see Fig. 1a).

We write  $\lambda_t^{r,f}$  to refer to the disjunction of the lineage expressions of the tuples in relation  $r$  with fact  $f$  that are valid at time point  $t$ . We write  $\lambda_t^{r,\theta}$  to refer to the disjunction of the lineage expressions of the tuples in relation  $r$  that satisfy  $\theta$  and are valid at time point  $t$ . When there are no tuples in  $r$  with fact  $f$  or satisfying  $\theta$  at time point  $t$ , we write  $\lambda_t^{r,f} = \text{null}$  or  $\lambda_t^{r,\theta} = \text{null}$ , respectively. We write  $\theta_{\tilde{r}}$  to indicate that values of attributes in condition  $\theta$  are instantiated to the corresponding values in tuple  $\tilde{r}$ . For example, for the  $\theta$  condition used in the query of Figure 1b and  $\tilde{r} = (\text{'Ann, ZAK, hotel}_1', a_1 \wedge b_3, [4, 6], 0.49)$ , we get  $\theta_{\tilde{r}} : b.\text{Loc} = \text{'ZAK'}$ .

The semantics of the TP data model are centered around two properties: TP snapshot reducibility and TP change preservation [5]. TP Snapshot reducibility states that the result of  $op^{\text{TP}}$  at each time point  $t$  is equal to the result of  $op^p$  on the input tuples with non-zero probability to be valid at  $t$ . Thus, the output attributes are determined only by the input tuples at  $t$  and the output lineages and probabilities are consistent with the possible-worlds semantics [14], [15]. The TP left outer join of Fig. 1b complies with TP snapshot-reducibility. For example, in tuple  $(\text{'Ann, ZAK, hotel}_1', [4,6], a_1 \wedge b_3, 0.42)$ , at time point  $t = 4$ , the fact is a combination of  $a_1.F = \text{'Ann, ZAK'}$  and  $b_3.F = \text{'hotel}_1, \text{'ZAK'}$ , i.e., the only input tuples valid at  $t$  and whose facts satisfy the join condition.

TP change preservation ensures that only consecutive time points of output tuples with equal facts and equivalent lineage expressions are grouped into intervals. It guarantees maximal intervals where the lineage expression is the same at all time points in the interval and different at time points outside. For example, the output tuples  $(\text{'Ann, ZAK, -'}, [2,4], a_1, 0.7)$  and  $(\text{'Ann, ZAK, -'}, [4,5], a_1 \wedge \neg b_3, 0.42)$  were not merged into the interval  $[2, 5)$ , since they do not have equivalent lineages.

#### IV. NEGATION IN TPDBS

The characterization of joins as operations with and without negation has been well established in databases [17]. As illustrated in Table I, the Cartesian product and the inner join are joins without negation since they only record information valid in both input relations. The anti join is a join purely based on negation and outer joins combine joins with and without negation.

TABLE I: Join Operations Categorized Based on Negation

	Operations
WITHOUT	$\times, \bowtie$
WITH	$\triangleright$
MIXED	$\bowtie\triangleright, \bowtie\bowtie, \bowtie\bowtie\bowtie$

A join with negation is performed over a positive relation  $p$  and a negative relation  $n$ . The result of a temporal-probabilistic join with negation includes, at each time point, the probability with which a tuple  $\tilde{p}$  of the positive relation  $p$  matches no tuple in the negative relation  $n$  under a predicate  $\theta$ . Firstly, this occurs at time points when either no tuple of  $n$  has non-zero probability to be valid or no valid tuple of  $n$  satisfies the  $\theta$ -condition. In this case, tuple  $\tilde{p}$  remains *unmatched* and the probability of the output tuple produced is equal to the probability of  $\tilde{p}$ .

Secondly, the non-existence of a matching tuple for  $\tilde{p}$  in  $n$  occurs when all the valid tuples of  $n$  that match  $\tilde{p}$  for  $\theta$  are false. This case relates to the probabilistic dimension and thus  $\tilde{p}$  is not disqualified for the output. The output fact is determined by  $\tilde{p}$  whereas for the computation of the corresponding probability we need to consider the negating form of the probabilities for the matching tuples in the negative relation. In case one of the matching tuples in  $n$  has probability equal to 1, the output tuple has 0 probability to be *true*.

*Example 2:* In Fig. 3, the TP anti join of relations  $a$  and  $b$  of Fig. 1a contains, at each time point, the probability that clients want to visit a location and no hotel is available. Tuple  $(\text{'Ann, ZAK'}, a_1, [2,4], 0.7)$  indicates that the tuple  $a_1$  of the positive relation  $a$  remains unmatched since there is no hotel in ZAK that has a probability to be available in the interval  $[2,4)$ . Tuple  $(\text{'Ann, ZAK'}, a_1 \wedge \neg(b_3 \vee b_2), [5,6], 0.084)$  corresponds to the case when the matching tuples of the negative relation  $b$  are *false*.

$Q = a \triangleright_{\theta}^{\text{TP}} b$				
Name	Loc	$\lambda$	$T$	$p$
Ann	ZAK	$a_1$	[2,4)	0.7
Ann	ZAK	$a_1 \wedge \neg b_3$	[4,5)	0.21
Ann	ZAK	$a_1 \wedge \neg(b_3 \vee b_2)$	[5,6)	0.084
Ann	ZAK	$a_1 \wedge \neg b_2$	[6,8)	0.28
Jim	WEN	$a_2$	[7,10)	0.8

Fig. 3:  $a \triangleright_{\theta}^{\text{TP}} b$  with  $\theta : a.\text{Loc} = b.\text{Loc}$  ( $a, b$  of Fig. 1a).

TP outer joins are joins with and without negation. What differs for outer joins when the temporal and the probabilistic dimension coexist is that two outcomes might arise at a time point. For example, in Fig. 1b, the TP left join  $a \bowtie^{\text{TP}} b$  includes, at each time point, cases when there is a non-zero probability for a tuple in  $a$  either to be matched with a tuple in  $b$  or not based on a predicate  $\theta$ . At time point  $t = 5$ , tuple  $a_1$  is combined with tuple  $b_3$  producing the output tuples  $(\text{'Ann, ZAK, hotel}_2', a_1 \wedge b_3, [4,6], 0.49)$  and  $(\text{'Ann, ZAK'}, -, a_1 \wedge \neg b_3), [4,5), 0.21)$  when  $b_3$  is *true* and *false*, respectively.

#### V. GENERALIZED WINDOWS

The use of a general  $\theta$  condition in TP outer joins and anti joins requires pairing input tuples that include different facts and combining multiple input tuples that are valid over an interval and satisfy  $\theta$ . For this purpose, we introduce *generalized lineage-aware temporal windows*, a mechanism created based on two TP relations  $r$  and  $s$ , with schema  $(F_T,$

TABLE II

Overlapping Windows	$\tilde{w} \in \mathbf{W}_O(\mathbf{r}; \mathbf{s}, \theta) \iff \exists r \in \mathbf{r}, s \in \mathbf{s} ( \tilde{w}.F_r = r.F \wedge \tilde{w}.F_s = s.F \wedge \theta \wedge \tilde{w}.\lambda_r \equiv r.\lambda \wedge \tilde{w}.\lambda_s \equiv s.\lambda \wedge \tilde{w}.T = r.T \cap s.T )$
Unmatched Windows	$\tilde{w} \in \mathbf{W}_U(\mathbf{r}; \mathbf{s}, \theta) \iff \tilde{w}.\lambda_s = \text{null} \wedge \tilde{w}.F_s = \text{null} \wedge \forall t \in \tilde{w}.T ( \exists r \in \mathbf{r} ( \tilde{w}.F_r = r.F \wedge \tilde{w}.\lambda_r \equiv r.\lambda ) \wedge \tilde{w}.\lambda_s \equiv \lambda_t^{\mathbf{s}, \theta_{\tilde{w}}} \wedge \lambda_t^{\mathbf{s}, \theta_{\tilde{w}}} = \text{null} ) \wedge \forall t' \notin \tilde{w}.T ( \nexists r \in \mathbf{r} ( \tilde{w}.F_r = r.F \wedge \tilde{w}.\lambda_r \equiv r.\lambda ) \vee \tilde{w}.\lambda_s \not\equiv \lambda_{t'}^{\mathbf{s}, \theta_{\tilde{w}}} )$
Negating Windows	$\tilde{w} \in \mathbf{W}_N(\mathbf{r}; \mathbf{s}, \theta) \iff \forall t \in \tilde{w}.T ( \exists r \in \mathbf{r} ( \tilde{w}.F_r = r.F \wedge \tilde{w}.\lambda_r \equiv r.\lambda ) \wedge \tilde{w}.F_s = \text{null} \wedge \lambda_t^{\mathbf{s}, \theta_{\tilde{w}}} \neq \text{null} \wedge \tilde{w}.\lambda_s = \lambda_t^{\mathbf{s}, \theta_{\tilde{w}}} ) \wedge \forall t' \notin \tilde{w}.T ( \nexists r \in \mathbf{r} ( \tilde{w}.F_r = r.F \wedge \tilde{w}.\lambda_r \equiv r.\lambda ) \vee \tilde{w}.\lambda_s \not\equiv \lambda_{t'}^{\mathbf{s}, \theta_{\tilde{w}}} )$

$F_r, T, \lambda_r, \lambda_s$ ).  $F_r$  and  $F_s$  are the facts included in tuples of relations  $\mathbf{r}$  and  $\mathbf{s}$  over interval  $T$ , respectively.  $\lambda_r$  is the disjunction of the lineage expressions of the tuples of relation  $\mathbf{r}$  that are valid over  $T$ , include  $F_r$  and satisfy  $\theta$ .  $\lambda_s$  is the disjunction of the lineage expressions of the tuples of relation  $\mathbf{s}$  that are valid over  $T$ , include  $F_s$  and satisfy  $\theta$ .

*Definition 1:* Let  $\mathbf{r}$  and  $\mathbf{s}$  be TP relations with schema  $(F, \lambda, T, p)$  and  $\theta$  a condition between the non-temporal attributes of  $\mathbf{r}$  and  $\mathbf{s}$ . The unmatched  $\mathbf{W}_U(\mathbf{r}; \mathbf{s}, \theta)$ , overlapping  $\mathbf{W}_O(\mathbf{r}; \mathbf{s}, \theta)$  and negating  $\mathbf{W}_N(\mathbf{r}; \mathbf{s}, \theta)$  windows of  $\mathbf{r}$  with respect to  $\mathbf{s}$  and  $\theta$  are defined according to Table II.

The *overlapping windows*  $\mathbf{W}_O(\mathbf{r}; \mathbf{s}, \theta)$  span a maximal interval over which a tuple  $r$  of  $\mathbf{r}$  overlaps with a tuple  $s$  from  $\mathbf{s}$  and the predicate  $\theta$  is satisfied. Tuple  $r$  includes the fact  $F_r$  and has lineage  $\lambda_r$  while  $F_s$  and  $\lambda_s$  correspond to the fact and lineage of tuple  $s$ . The interval of the window that is produced by the pair of tuples  $r$  and  $s$  corresponds to the overlap of their interval ( $\tilde{w}.T = r.T \cap s.T$ ). The *unmatched windows*  $\mathbf{W}_U(\mathbf{r}; \mathbf{s}, \theta)$  span over the interval or a subinterval of a tuple  $r$  of  $\mathbf{r}$  during which all tuples of  $\mathbf{s}$  are either not valid or don't satisfy  $\theta$  ( $\lambda_t^{\mathbf{s}, \theta_{\tilde{w}}} = \text{null}$ ). The fact  $F_r$  and the lineage  $\lambda_r$  of an unmatched window are determined by  $r$  while  $F_s$  and  $\lambda_s$  are set to null. The negating windows  $\mathbf{W}_N(\mathbf{r}; \mathbf{s}, \theta)$  of the TP relation  $\mathbf{r}$  with respect to the TP relation  $\mathbf{s}$  are windows during which a fact is included in a tuple  $r$  of  $\mathbf{r}$  as well as in multiple tuples of  $\mathbf{s}$  that are valid and satisfy the  $\theta$ -condition. Negating windows are suitable for producing output tuples where, for  $\theta$ , all the tuples of  $\mathbf{s}$  that match a tuple  $r$  of  $\mathbf{r}$  including the fact  $F_r$  are *false*, as described in Section IV. Thus, the fact  $F_r$  and the lineage  $\lambda_r$  of the window are determined by  $r$ ,  $F_s$  is set to null and  $\lambda_s$  is the disjunction of the lineages of all the tuples in  $\mathbf{s}$  that match  $r$ .

*Example 3:* In Fig. 4, the TP relations  $\mathbf{a}$  and  $\mathbf{b}$  of Fig. 1 are illustrated along with the unmatched, overlapping and negating windows of  $\mathbf{a}$  with respect to  $\mathbf{b}$ . Single lines are used for tuples. Pairs of lines denote windows. Different colors are used to annotate different facts: black is used for 'Ann, ZAK', red for 'John, WEN', green for 'hotel<sub>3</sub>, SOR', yellow for 'hotel<sub>2</sub>, ZAK', and blue for 'hotel<sub>1</sub>, ZAK'. Wavy lines are used for tuples of an input relation that match no tuple of the other relation for  $\theta$ . For the unmatched window  $w_1 = (\text{'Ann, ZAK', null}, [2, 4], a_1, \text{null})$ , the straight black line indicates that the fact  $w_1.F_r = \text{'Ann, ZAK'}$  and the lineage  $w_1.\lambda_r = a_1$  match the corresponding attributes of tuple  $a_1$ . The dotted line indicates that fact  $w_1.F_s$  is null and so is  $w_1.\lambda_s$ . At  $t = 4$ ,  $a_1$  is

still valid whereas  $\lambda_4^{\mathbf{b}, \theta_{w_1}} = b_3$ , which indicates that a tuple of  $\mathbf{b}$  starts being valid and thus interval  $[2, 4)$  is maximal. The window  $w_3 = (\text{'Ann, ZAK', 'hotel<sub>1</sub>'}, [4, 6], a_1, b_3)$  is an overlapping window. The blue and a black straight lines for  $w_1$  indicate that  $F_r$  and  $F_s$  of  $w_3$  correspond to the facts of tuples  $a_1$  and  $b_3$ , i.e., tuples that overlap and include the same values for *Loc*. For the negating window  $w_6 = (\text{'Ann, ZAK', null}, [5, 6], a_1, b_3 \vee b_2)$ , the black straight line in  $w_6$  indicates that its fact  $F_r$  and its lineage  $\lambda_r$  correspond to the fact and lineage of  $a_1$ . The fact  $F_s$  is null, illustrated by a dotted line. Annotated next to this line, the  $\lambda_s$  equals the disjunction of the tuples  $b_2$  and  $b_3$  that satisfy  $\theta$  over the interval  $[5, 6)$ . The interval  $[5, 6)$  is maximal since at  $t = 6$ ,  $b_3$  stops being valid.

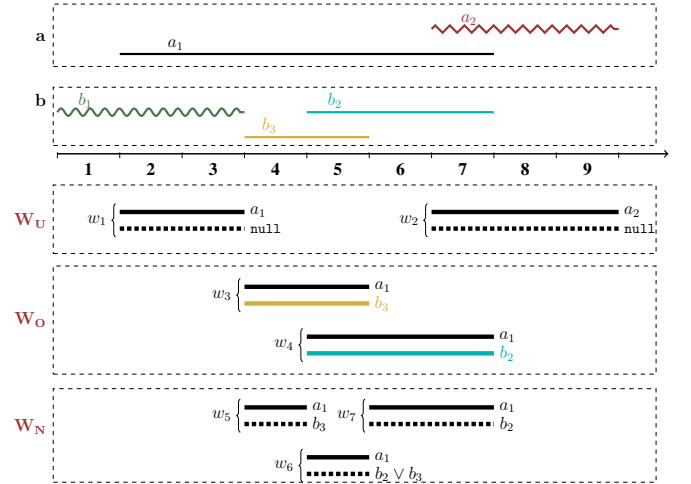


Fig. 4: All windows of  $\mathbf{a}$  with respect to  $\mathbf{b}$  with  $\theta : \mathbf{a}.Loc = \mathbf{b}.Loc$

An output tuple is formed for each window using the facts ( $F_r, F_s$ ) and interval  $T$  in their exact form while the output lineage is formed by combining  $\lambda_r$  and  $\lambda_s$  with the proper lineage-concatenation function. According to their semantics, each set of windows is matched with a unique function: for *overlapping windows* we use the function **and**, for *negating windows* we use **andNot** and for *unmatched windows* only  $\lambda_r$  is passed on to the output lineage. For the TP anti join in Figure 3, the unmatched window  $(\text{'Ann, ZAK', null}, [2, 4], a_1, \text{null})$  is transformed to the output tuple  $(\text{'Ann, ZAK', -}, [2, 4], a_1)$  and the negating window  $(\text{'Ann, ZAK', null}, [5, 6], a_1, b_3 \vee b_2)$  is transformed to the output tuple  $(\text{'Ann, ZAK', -}, [5, 6], a_1 \wedge \neg(b_3 \vee b_2))$ . In Table III, we include all the window sets required for each TP join with negation considering that

$$W_o(r; s, \theta) = W_o(s; r, \theta).$$

TABLE III: TP Joins with Negation using Windows

$op^{TP}$	$W_U(r; s, \theta)$	$W_N(r; s, \theta)$	$W_o(r; s, \theta)$	$W_U(s; r, \theta)$	$W_N(s; r, \theta)$
$r \supset s$	✓	✓			
$r \sqsupset s$	✓	✓	✓		
$r \sqsubset s$			✓	✓	✓
$r \sqsubset s$	✓	✓	✓	✓	✓

## VI. ALGORITHMS

In this section, we introduce algorithms to compute *generalized lineage-aware temporal windows* and the result of TP joins with negation. Our Lineage-Aware Window Advancers (LAWA) for unmatched (LAWA<sub>U</sub>) and negating (LAWA<sub>N</sub>) windows use overlapping windows as a computational basis. LAWA<sub>U</sub> (Algorithm 1) produces the unmatched windows of  $r$  with respect to  $s$  by identifying the subintervals of  $r$  during which there is no overlap or match with a tuple of  $s$ , i.e., subintervals that do not correspond to any overlapping window. Similarly, each of the negating windows of  $r$  with respect to  $s$  spans a subinterval where all tuples of  $s$  that overlap and match with a tuple  $r$  of  $r$  are false and thus lineage information from all the overlapping windows that are valid over this subinterval and involving  $r$  must be combined.

LAWA<sub>U</sub> and LAWA<sub>N</sub> are sweeping-window algorithms [5] that are applied on windows instead of tuples. They are responsible for forming a set of windows based on overlapping ones but also for passing the input windows to the output since they are also necessary for the result of a TP join with negation. They are operating in an incremental manner, thus avoiding recomputing the overlapping windows multiple times.

### A. Overlapping Windows

For the computation of overlapping windows of relation  $r$  with respect to  $s$ , we perform the conventional outer join  $r \bowtie_{\theta_o \wedge \theta} s$  with the overlapping predicate  $\theta_o : r.T \cap s.T$  and a condition  $\theta$  on the non-temporal attributes, as provided in the TP join to be computed. The result of  $r \bowtie_{\theta_o \wedge \theta} s$  computes a set of windows enhanced with the time-interval of the tuple of  $r$  valid over each window, and its result has schema:  $(F_r, \lambda_r, F_s, \lambda_s, [O_s, O_e], [T_s, T_e])$ .  $(F_r, [T_s, T_e], \lambda_r)$  correspond to the fact, interval and lineage of a tuple  $r$  in  $r$ . Similarly,  $(F_s, \lambda_s)$  correspond to tuple  $s$  in  $s$ .  $[O_s, O_e]$  is the interval during which the tuples  $r$  and  $s$  overlap.

	X					
	$F_r$	$\lambda_r$	$F_s$	$\lambda_s$	$[O_s, O_e]$	$[T_s, T_e]$
$x_1$	'Ann, ZAK'	$a_1$	'hotel1, ZAK'	$b_3$	[4,6)	[2,8)
$x_2$	'Ann, ZAK'	$a_1$	'hotel2, ZAK'	$b_2$	[5,8)	[2,8)
$x_3$	'Jim, WEN'	$a_2$	null	null	null	[9,12)

Fig. 5: The result of  $a \bowtie_{L_T \cap s.T \wedge a.Loc=b.Loc} b$ .

The tuples of the join  $r \bowtie_{\theta_o \wedge \theta} s$  for which all attributes are not null constitute the set of overlapping windows  $W_o(r; s, \theta)$ . However, the use of the conventional left join results also in pairs with null attributes.

### B. Unmatched Windows

The unmatched windows of a TP relation  $r$  with respect to a TP relation  $s$  and a condition  $\theta$  are computed in two phases. Firstly, the windows in result of  $r \bowtie_{\theta_o \wedge \theta} s$  with  $(F_s, \lambda_s)$  and  $[O_s, O_e]$  equal to null correspond to unmatched windows where input tuples of  $r$  don't overlap or satisfy  $\theta$  with any tuple in  $s$ . The interval of each such window is equal to the interval  $[T_s, T_e]$  of the tuple of  $r$ .

Secondly, the algorithm LAWA<sub>U</sub> extends the result  $X$  of  $r \bowtie_{\theta_o \wedge \theta} s$  (cf. Fig. 5) with the remaining unmatched windows, i.e., the windows that span a *subinterval* of a tuple in  $r$  during which no tuple in  $s$  is valid or satisfies  $\theta$ . For these unmatched windows to be created, the windows in  $X$  are grouped according to the fact  $F_r$  and the interval  $[T_s, T_e]$  of the tuple in  $r$  to which they correspond. Within each group, the tuples are sorted on the starting point ( $O_s$ ) of the overlapping intervals and the order of tuples with equal starting points does not matter. The algorithm performs a sweep of the interval  $[T_s, T_e]$  of each  $r$  tuple of  $r$ . It copies the overlapping windows ( $[O_s, O_e] \neq \text{null}$ ) relating to  $r$  to the output. At the same time, given the subintervals that the overlapping windows span and the initial interval  $[T_s, T_e]$  of  $r$ , it identifies the subintervals during which there is no overlap with a tuple in  $s$ , i.e., no overlapping window, and produces the remaining unmatched windows.

#### Algorithm 1: LAWA<sub>U</sub>(status)

```

1 (prevWindTe,  $F_r$ ,  $\lambda_r$ , wind, PQ, neg) = status;
2 if wind = null then return null;
3 do
4   if prevWindTe = -1 then
5     | windTs = wind. $T_s$ ;  $F_r$  = wind. $F_r$ ;  $\lambda_r$  = wind. $\lambda_r$ ;
6   else windTs = prevWindTe;
7    $\lambda_s$  = null;  $F_s$  = null;
8   if wind. $O_s$  = windTs then
9     |  $\lambda_s$  = wind. $\lambda_s$ ;  $F_s$  = wind. $F_s$ ;
10  if  $\lambda_s \neq \text{null}$  then windTe = wind. $O_e$ ; // Case 1
11  else if windTs = wind. $T_s$   $\wedge$  wind. $O_s \neq \text{null}$  then
12    | windTe = wind. $O_s$ ; // Case 2
13  else if wind. $O_s$  = null  $\vee$  windTs = wind. $O_e$  then
14    next = getNextOf(wind);
15    if next  $\neq \text{null}$   $\wedge$   $F_r$  = next. $F_r$ ;
16    then // Case 3
17      | windTe = next. $O_s$ ;
18    else windTe = wind. $T_e$ ; // Case 4, 5
19    wind = next;
20  if windTe = wind. $T_e$  then prevWindTe = -1;
21  else prevWindTe = windTe;
22 while windTs  $\geq$  windTe;
23 out = ( $F_r$ ,  $F_s$ , windTs, windTe,  $\lambda_r$ ,  $\lambda_s$ );
24 status = (prevWindTe,  $F_r$ ,  $\lambda_r$ , wind, PQ, neg);
25 return (window, status);

```

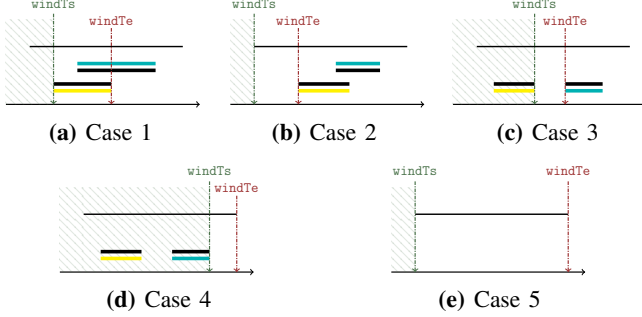
The execution of algorithms LAWA<sub>U</sub> and LAWA<sub>N</sub> is based on a context node (status) with information on the status of



the algorithm: the right boundary of the last output window (prevWindTe), the fact ( $F_r$ ) and the lineage ( $\lambda_r$ ) of the tuple of  $r$  that is valid over the output window [windTs, windTe), and the input window (wind) to be processed. The tag neg and the priority queue PQ are not used in  $LAWA_U$ . At each call, a generalized lineage-aware temporal window out is returned (Line 23) as well as the status necessary for the next call. Prior to the first call of  $LAWA_U$ , the first window of  $X$  is fetched,  $F_r$  and  $\lambda_r$  are initialized to null and prevWindTe is initialized to  $-1$ .

**Lines 4-6:** Initially, the left boundary windTs of the new window as well as the fact and the lineage of the valid tuple of  $r$  are determined. If a new group is being processed (prevWindTe =  $-1$ ), windTs is determined by the starting point of the first window wind of the new group. In this case, the fact  $F_r$  and the lineage  $\lambda_r$  of the valid tuple of  $r$  are also extracted from wind. If the processing of a group continues, the interval of the new window is adjacent to the previous one, with windTs = prevWindTe while  $F_r$  and  $\lambda_r$  remain unchanged.

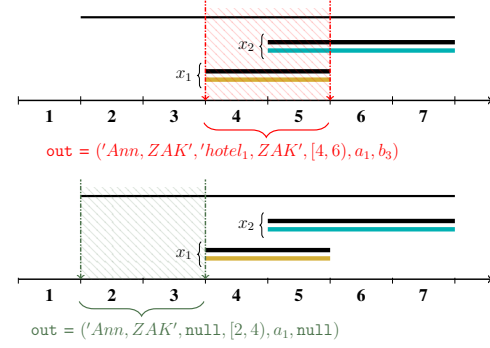
**Lines 7-9:** In order to determine the fact and the lineage of the tuple of  $s$  valid over the output window, we check if the starting point windTs of the window matches the starting point  $O_s$  of an overlapping window in  $X$ . If satisfied, this condition (Line 8) indicates that there is a tuple of  $s$  valid over the window and thus the fact  $F_s$  and lineage  $\lambda_s$  equal the corresponding attributes of wind. Otherwise, they are set to null.



**Fig. 6:** Cases for determining windTe in  $LAWA_U$  Algorithm. Single line is used for the input tuple and pairs of lines for the windows.

**Lines 10-19:** The right boundary windTe of out is determined based on whether it is an overlapping or an unmatched one. All the cases are annotated in the algorithm and illustrated in Figure 6. If out is an overlapping window (Case 1), i.e.,  $\lambda_s \neq \text{null}$ , its interval corresponds to the overlapping interval in wind and thus, windTe is set to wind.Oe. If the output window is an unmatched window, three different cases are considered based on the position of windTs with respect to [wind.Os, wind.Oe). If the starting point windTs coincides with the starting point of the valid tuple of  $r$  (windTs = wind.Ts) and the starting point of the overlapping window wind succeeds (Case 2), windTe is set to the starting point of wind. If the starting point of the output window coincides with the ending point of the overlapping window

wind (Case 3), the upcoming window next is fetched. If next is in the same group as wind, out is positioned between two overlapping windows and thus windTe = next.Os. However, if next belongs to a new group, wind is positioned at the end of the interval of a valid tuple of  $r$  (Case 4). Thus windTe = wind.Te and the sweeping progresses to window next. The same assignment takes place if wind is one of the unmatched windows produced by the conventional left outer join (Case 5).



**Fig. 7:**  $LAWA_U$  on the group with  $F_L = 'Ann, ZAK'$  and  $\lambda_L = a_1$ .

**Example 4:** In Fig. 7, we illustrate two calls of  $LAWA_U$  when applied on relation  $X$  of Fig 5 and more specifically on the group of windows with the fact  $F_r = 'Ann, ZAK'$ . The single blank line corresponds to tuple  $a_1$ , the tuple of the left relation a valid over all windows of the group. The window wind =  $x_1$  is the first to be processed. In the first call of  $LAWA_U$ , illustrated at the bottom of the figure, the processing of a new group starts and windTs,  $F_r$  and  $\lambda_r$  are initialized to the starting point, fact and lineage of  $a_1$ , respectively. No overlapping window of the same group starts at windTs = 2 and thus,  $F_s$  and  $\lambda_s$  are set to null. According to Case 2, windTe is set to wind.Os. In the second call of  $LAWA_U$ , the same group is processed and out will be adjacent to the previous output window. Since windTs equals the starting point of the overlapping window  $x_1$ , the facts, lineages and intervals of the output window are fetched from  $x_1$ . The ending point windTe of out is set according to Case 1.

### C. Negating Windows

$LAWA_N$  extends the result  $Y$  of  $LAWA_U$  with the negating windows.  $Y$  consists of windows ordered by the fact of  $r$  ( $F_r$ ) as well as by their starting point ( $T_s$ ).  $LAWA_N$  sweeps over  $Y$  and copies all the unmatched and overlapping windows to the output. When a group of overlapping windows with the same fact  $F_r$  is encountered, negating windows are created. The intervals of these windows are subintervals of the group of overlapping windows.

The execution of  $LAWA_N$  is also based on the context node status. The tag neg indicates if a negating window will be produced. The priority queue PQ includes  $(t, \lambda)$  pairs that indicate the time point  $t$  after which the tuple of the right relation with lineage  $\lambda$  stops being valid.

Y				
	$F_r$	$F_s$	$\lambda_r \quad \lambda_s$	$T = [T_s, T_e]$
y <sub>1</sub>	'Ann, ZAK'	null	a <sub>1</sub> null	[2,4)
y <sub>2</sub>	'Ann, ZAK'	'hotel <sub>1</sub> , ZAK'	a <sub>1</sub> b <sub>3</sub>	[4,6)
y <sub>3</sub>	'Ann, ZAK'	'hotel <sub>1</sub> , ZAK'	a <sub>1</sub> b <sub>2</sub>	[5,8)
y <sub>4</sub>	'Jim, WEN'	null	a <sub>2</sub> null	[9,12)

Fig. 8: The input of  $LAWA_N$

**Lines 1-6:** In the first call of the algorithm (`firstCall`), the first tuple of  $Y$  is fetched, the priority queue PQ is initialized (pointer to null), `prevWindTe` is set to  $-1$  and `neg` to *false*. Since negating windows are created based on the overlapping windows, whenever a group of overlapping windows with the same  $F_r$  starts, the output fact  $F_r$ , the output lineage  $\lambda_r$  and the starting point `prevWindTe` of the output windows are updated to the values of the first tuple of this group for  $F_r$ ,  $\lambda_r$  and  $T_s$  respectively.

---

**Algorithm 2:**  $LAWA_N(\text{status})$

---

```

1 (prevWindTe,  $F_r$ ,  $\lambda_r$ , wind, PQ, neg) = status;
2 if wind = null  $\wedge$  isPQempty() then return (null, null);
3 if firstCall then
4   PQ = initializePQ(); prevWindTe = -1; neg = false;
5 if prevWindTe = -1  $\wedge$  wind. $\lambda_r \neq$  null then
6    $F_r$  = wind. $F_r$ ;  $\lambda_r$  = wind. $\lambda_r$ ; prevWindTe = wind. $T_s$ ;
7 while out = null do
8   if neg = false then
9     out = wind;
10    if wind. $F_s$  = null then wind = getNextTuple();
11    else neg = true; addToPQ(wind. $T_e$ , wind. $\lambda_s$ );
12  else if wind. $F_r$  =  $F_r \wedge$  wind. $T_s \leq$  prevWindTe then
13    wind = getNextTuple();
14  if out = null  $\wedge$  wind. $F_r$  =  $F_r$  then
15    if wind. $T_s >$  prevWindTe then
16      windTe = tForTopOfPQ();
17      if wind. $T_s <$  windTe then
18        windTe = wind. $T_s$ ;
19         $\lambda_s$  = disjunctLineages(windTe);
20        out = ( $F_r$ , -, [prevWindTe, windTe),  $\lambda_r$ ,  $\lambda_s$ );
21        prevWindTe = windTe;
22        neg = false;
23    else if wind. $T_s$  = prevWindTe then neg = false;
24  else if out = null  $\wedge$  ( $\neg$  isPQempty()) then
25    windTe = tForTopOfPQ();
26     $\lambda_s$  = disjunctLineages(windTe);
27    out = ( $F_r$ , -, [prevWindTe, windTe),  $\lambda_r$ ,  $\lambda_s$ );
28    prevWindTe = windTe; removeTopOfPQ();
29 if isPQempty() then prevWindTe = -1; neg = false;
30 status = (prevWindTe,  $F_r$ ,  $\lambda_r$ , wind, PQ, neg);
31 return (out, status);

```

---

**Lines 8-13:**  $LAWA_N$  outputs an unmatched, overlapping or negating window according to `neg`. When `neg` is *false* (Line 8), the unmatched or overlapping window `wind` is copied to the output as is (Line `refline:copy`). If `wind` corresponds to an unmatched window (wind. $F_s$  = null), we proceed to

the next window. However, if it corresponds to an overlapping window, the creation of a negating window follows and `neg` is set to *true* (Line 11). In this case, we add to PQ the pair (wind. $T_e$ , wind. $\lambda_s$ ), with the ending point and the lineage of the valid tuple in the relation  $s$  as recorded in `wind`.

When `neg` is *true*, the creation of a negating window follows. If the same group is processed and the starting point of `out` (`prevWindTe`) is equal to the starting point of `wind`, the next window is fetched (Line 13) for two reasons. Firstly, if the next window of  $Y$  is an overlapping window of the same group and starts at `prevWindTe`, the lineage of the tuple of relation  $s$  valid over this input window needs to be considered for  $\lambda_s$ . Secondly, if the next window belongs to the same group, its starting point should be considered as a potential ending point of `out`.

**Lines 14-23:** The output negating window is finalized by determining its ending point `windTe` and lineage  $\lambda_s$ . The lineage  $\lambda_s$  is always determined by disjuncting the lineage expressions of the pairs  $(t, \lambda)$  in the priority queue with  $t$  smaller than `windTe`. Thus,  $\lambda_s$  correspond to the disjunction of the tuples of the relation  $s$  valid over the output interval [`prevWindTe`, `windTe`). To determine the ending point `windTe` of the window, we first check if the upcoming window `wind` of  $Y$  includes the same fact  $F_r$  as `out`. If this is the case, `windTe` is the minimum between the time point of the top pair in the queue, i.e., the smallest ending point of valid tuples in relation  $s$ , and the starting point of the upcoming window of  $Y$ . Therefore, a window is created when there is a change in the tuples of relation  $s$  that are valid either because a tuple ends or a new tuple begins. After `out` is formed, the starting point `prevWindTe` of the next negating window is set to `windTe`. `neg` is set to *false* so that the window `wind` is copied to the output.

A special case occurs when the starting point of the upcoming window is equal to the starting point of the output window (Line 23). This means that there exists a valid tuple in the reference relation  $s$  that needs to be considered for the output window and thus its finalization is postponed. The upcoming window, either overlapping or unmatched, has to be first copied to the output so we set `neg` back to *false*.

**Lines 24-28:** If there are more overlapping windows in PQ that end before the upcoming window `wind` starts, regardless of whether `wind` belongs in the same or a different group, the ending point of the new negating window is equal to the ending point of the pair on top of the priority queue (Line 25). The starting point of the next negating window is set to `windTe` indicating that the sweeping until this time point has been completed. As a result, all the nodes in PQ correspond to windows whose ending point is equal to `windTe` have already been considered and need to be removed.

**Example 5:** In Fig. 9, we focus on the group with  $F_r$ ='Ann, ZAK' and we illustrate all six calls of  $LAWA_N$  on the corresponding windows of the result  $Y$  of  $LAWA_U$  (Fig.8), when applied on the relations  $a$  and  $b$  of Fig.1a. Red color is used for windows copied to the output whereas green is used for the negating windows. In the first two calls of  $LAWA_N$ ,



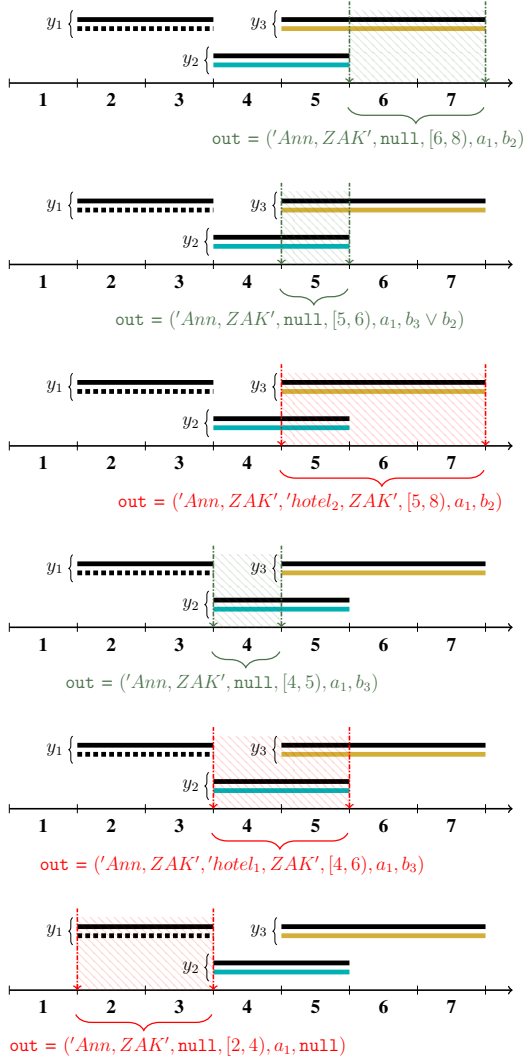


Fig. 9: Execution of  $LAWA_N$  on the result of  $LAWA_U$

windows  $y_1$  and  $y_2$  are copied to the output.  $y_2$  is the first overlapping window after a series of unmatched ones. After  $out = y_2$ ,  $neg$  is set to *true* and the sweeping for negating tuples starts from  $prevWindTe = y_2.Ts = 4$  with  $F_\lambda = 'Ann, ZAK'$  and  $\lambda_r = a_1$ . Window  $y_2$  is followed by another overlapping window ( $y_3$ ) that starts before the ending point of  $y_2$ , recorded in the top node of the priority queue. Consequently,  $windTe = y_4.Ts = 5$  and the negating window  $('Ann, ZAK', null, [4, 5], a_1, b_3)$  is produced.  $neg$  is set false and window  $y_3$  is then copied to the output. Since there are no more overlapping windows to be processed, the upcoming negating windows are adjacent to each other and their ending points are derived from the nodes of PQ.

#### D. TP Join Algorithms

In this subsection we introduce the algorithm  $NegationJoins(r, s, \theta, op)$  that computes the result of the TP outer join or anti join  $op$  on the input TP relations  $r$  and  $s$  and the predicate  $\theta$ . In contrast to previous works in either temporal

or probabilistic databases, this algorithm involves no tuple replication. Instead, it allows for a pipelined calculation of the result and thus enables its smooth integration in the kernel of a DBMS.

#### Algorithm 3: $NegationJoins(r, s, \theta, op)$

```

1  $w_{init} = leftJoin(r, s, \theta \wedge \theta_o)$ ;
2  $sort(w_{init} \{F_L, 0_s\})$ ;
3  $status = (-1, null, null, fetchWind(w_{init}), null, false)$ ;
4 while  $status \neq null$  do
5    $(w, status) = LAWA_u(status)$ ;
6    $w_{uo} = w_{uo} \cup \{w\}$ ;
7  $status = (-1, null, null, fetchWind(w_{uo}), null, false)$ ;
8 while  $status \neq null$  do
9    $(w, status) = LAWA_n(status)$ ;
10  if  $w.\lambda_s = null \wedge w.F_s = null$  then
11     $o = o \cup \{(w.F_r, w.F_s, w.\lambda_r, [w.winTs, w.winTe])\}$ ;
12  else if  $w.\lambda_s \neq null \wedge w.F_s = null$  then
13     $\lambda = andNot(w.\lambda_r, w.\lambda_s)$ ;
14     $o = o \cup \{(w.F_r, w.F_s, \lambda, [w.winTs, w.winTe])\}$ ;
15  else if  $op \neq \triangleright$  then
16     $\lambda = and(w.\lambda_r, w.\lambda_s)$ ;
17     $o = o \cup \{(w.F_r, w.F_s, \lambda, [w.winTs, w.winTe])\}$ ;
18 if  $op = \bowtie$  then  $o = o \cup NegatingJoins(s, r, \theta, \triangleright)$ ;
19 return  $o$ ;

```

Initially, the set  $w_{init}$  includes the overlapping windows of  $r$  and  $s$  and a subset of the unmatched windows (Section VI-A). The windows in  $w_{init}$  are sorted based on the fact  $F_r$  and the starting point  $Ts$  (Line 2) of the tuple of the positive relation from which they have been produced. As long as the terminating condition (Line 4) is satisfied,  $LAWA_u$  passes through all start and end points of the windows in  $w_{init}$  in a smaller-to-larger fashion and expands the set with the unmatched windows (Line 6) that hadn't been created yet. Similarly,  $LAWA_n$  sweeps the windows of the set  $w_{uo}$  and extends it with the negating windows of  $r$  and  $s$ .

Each window  $w$  that  $LAWA_n$  produces is not further swept and it can be transformed to an output tuple for the result of the TP join. A lineage-based filter is directly applied to determine if  $w$  is unmatched ( $w.\lambda_s = null \wedge w.F_s = null$ ), negating ( $w.\lambda_s \neq null \wedge w.F_s = null$ ) or overlapping. If the join performed is a TP anti join ( $\triangleright^{TP}$ ), then the overlapping windows are filtered out and are not included in the final result. If it is a full outer join, the unmatched and negating windows of  $s$  using  $r$  as a reference need to be included and thus the  $NegationJoins$  algorithm needs to be called again with reversed arguments, same predicate and anti join as the operation to be performed so that the overlapping windows are not copied again to the output. Finally, every window is finalized into an output tuple using the lineage-concatenating function that corresponds to set of windows to which it belongs. In the case of a TP anti join,  $F_r$  is the only fact included in the output tuples.

## VII. EVALUATION

In this section, we evaluate our algorithms using two real-world datasets which vary on (i) the number of facts in the input relations and (ii) the percentage of tuples whose intervals overlap. We compare our approach for TP joins with negation (NJ) to Temporal Alignment (TA), i.e., the only related approach that can be used for the computation of TP outer joins and TP anti join. The experiments show that our approach outperforms TA and it is the only scalable solution for TP joins with negation on input relations of more than 200K tuples. *NJ* is also robust with predictable performance with respect to the aforementioned characteristics of the datasets.

### A. Experimental Setup

All of the following experiments were deployed on a 2xIntel(R) Xeon(R) CPU E5-24400 @2.40GHz machine with 64GB main memory, running CentOS 6.7. Our algorithms have been implemented in the kernel of PostgreSQL in C, and all experiments were performed in main-memory. No indexes were used. In all PostgreSQL implementations, the maximum memory for sorting as well as for shared buffers were set to 10GB.

We have implemented *NJ* in PostgreSQL 9.4.3 by modifying the parser, executor and optimizer. The only approach our implementation can be compared against is **Temporal Alignment (TA)** [11]. **Temporal Alignment** is an approach developed for the computation of temporal operations using sequenced semantics and is implemented in the kernel of PostgreSQL as well. It consists of a set of reduction rules based on *Normalize* ( $\mathcal{N}$ ) and *Align* ( $\Phi$ ), two operators responsible for the interval adjustment of the input relations. Due to the existence of probabilities, the results of TP joins with negation differ and thus, for our experiments, we introduced reduction rules that are consistent with the TP semantics while properly exploiting  $\mathcal{N}$  and  $\Phi$ . For a fair comparison, we migrated the authors' implementation to PostgreSQL 9.4.3.

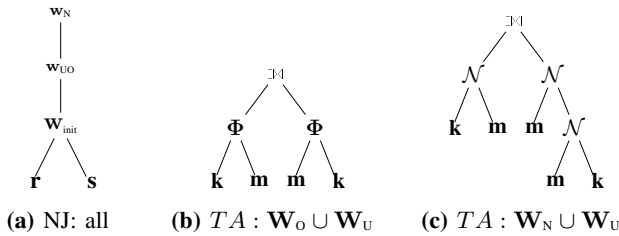


Fig. 10: Query Trees

In Fig. 10, we illustrate the query plans used by NJ and TA for the computation of windows. In Fig. 10a, the nodes  $w_{init}$ ,  $w_{uo}$  in the tree correspond to sets of windows as described in Algorithm 3. The node  $w_N$  corresponds to the set of negating windows produced by the calls of  $LAWA_N$ . In Fig. 10b and 10c, we illustrate the two query subtrees in TA for the computation of all output tuples. The operators  $\mathcal{N}$  and  $\Phi$  in TA replicate the tuples of the left relation and assign new intervals

based on the right relation. Since the facts and lineages of the input tuples still need to be combined, additional joins are performed.  $\Phi(k, m)$  is associated with overlapping windows (Fig. 10b) since the subintervals it produces correspond to the overlap of a tuple in  $k$  with a tuple in  $m$ .  $\mathcal{N}(k, m)$  is appropriate for negating windows since it includes intervals that correspond to the overlap of a tuple in  $k$  with a group of tuples in  $m$ . Both  $\Phi(k, m)$  and  $\mathcal{N}(k, m)$  include intervals where a tuple  $k$  in  $k$  matches no tuple in  $m$ , leading to the unmatched windows being computed twice. In Fig. 10c, the tuples of the right relation  $m$  are adjusted both using relation  $k$  and itself because, over an interval, we compute the tuples of  $m$  that are valid and are combined with a tuple of  $k$ . Given that  $\mathcal{N}$  only uses one input relation as reference, we need to further adjust  $m$  based on the result of  $\mathcal{N}(k, m)$ .

The  $\Join_{\theta \wedge \theta_o}$ ,  $\mathcal{N}$  and  $\Phi$  nodes are all based on a conventional left-outer join with a condition for the interval overlap of the matching tuples. PostgreSQL's optimizer determines whether such a join is executed as a nested loop, a merge join or a hash join depending on the  $\theta$  condition of the TP join to be computed.  $\Join_{\theta \wedge \theta_o}$  is computed using a nested loop only when the  $\theta$  condition used has low selectivity, i.e., when a high percentage of pairs of input tuples satisfy the condition. On the contrary, this varies for  $\mathcal{N}$  and  $\Phi$ , based on whether a TP join or a set of windows is computed.

### B. Real-World Datasets

The Webkit dataset<sup>1</sup> [27], [12], [28] records the history of 484K files of the SVN repository of the Webkit project over a period of 11 years at a granularity of milliseconds. Each tuple has schema  $(File\_Path, [T_s, T_e])$  and the valid times indicate the periods when a file remained unchanged. The Meteo Swiss dataset<sup>2</sup> includes temperature predictions that have been extracted from the website of the Swiss Federal Office of Meteorology and Climatology. Each tuple has schema  $(Station\_ID, Value\_ID, Value, [T_s, T_e])$ . The measurements were taken at 80 different meteorological stations ( $Station\_ID$ ) in Switzerland from 2005 to 2015 and involve four different metrics ( $Value\_ID$ ), including temperature and precipitation. Measurements are 10 minutes apart and – in order to produce intervals – we merged time points whose measurements differ by less than 0.1.

The main properties of these datasets are summarized in Table IV. For both datasets we produced a second relation by shifting the intervals of the original dataset, without modifying the lengths of the intervals. The start/end points of the new relation were chosen according to the distribution of the original ones.

### C. Runtime

In Fig. 11, 12, 13 we illustrate the runtime for the overlapping and unmatched windows, negating windows, and for a TP left outer join, respectively, over subsets of the Webkit and Meteo dataset. The subsets range from 20K to 200K tuples.

<sup>1</sup>The WebKit Open Source Project: <http://www.webkit.org> (2012)

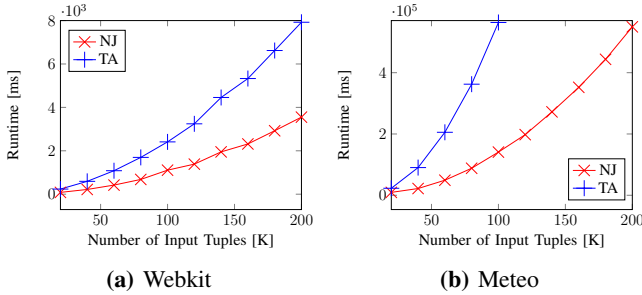
<sup>2</sup>Federal Office of Meteorology and Climatology: <http://www.meteoswiss.ch> (2016)

**TABLE IV: Real-World Dataset Properties**

	Meteo	Webkit
Cardinality	10.2M	1.5M
Time Range	347M	7M
Min. Duration	600	0.02
Max. Duration	19.3M	6M
Avg. Duration	152M	1.7M
Num. of Facts	80	484K
Distinct Points	545K	144K
Max Num. of Tuples (per time point)	140	369K
Avg Num. of Tuples (per time point)	37	21

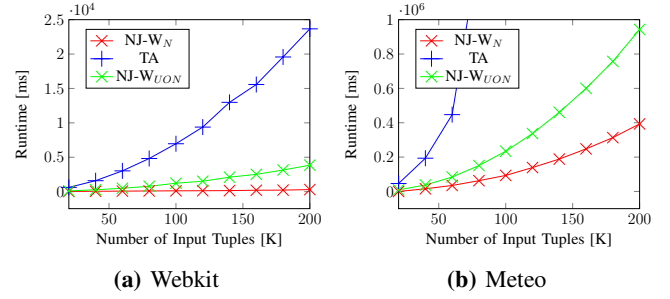
For Webkit dataset, as a  $\theta$  condition we apply equality of the *File\_Path*, i.e., we combine tuples referring to the same file. For Meteo dataset, we apply equality on *Value\_IDs* and inequality on *Station\_IDs*, i.e. we combine tuples with measurements on the same metric but taken in different stations.

Fig. 11 shows the runtime of NJ and TA for the set  $w_{uo}$  (Algorithm 3), including the unmatched and overlapping windows. Both approaches follow a similar trend and the reason is that the most computationally demanding part of both is a conventional left join, used to identify the pairs of tuples that overlap. As shown in Fig. 10, NJ only executes this join once whereas TA executes it twice. As a result, NJ is two to four times faster.

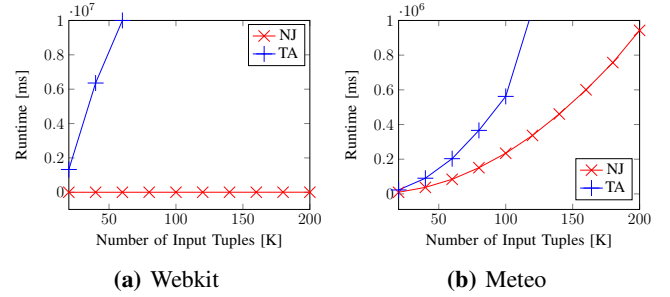

**Fig. 11:  $w_{uo}$ : Overlapping and Unmatched Windows**

In Fig. 12, we have illustrated the runtime for the computation of negating windows. In NJ, negating windows are computed by applying  $LAW_{A_N}$  on the set  $w_{uo}$ . Thus, we have illustrated their computation time both including ( $w_{uon}$ ) and excluding ( $w_N$ ) the runtime for  $w_{uo}$ . In the case of  $w_{uon}$ , NJ computes the negating windows four to ten times faster than TA whereas, in the case of  $w_N$ , it computes them twelve to twenty times faster.

Finally, the runtimes of both NJ and TA for a TP left-outer join are illustrated in Fig. 13. To compute the join with TA, a duplicate-eliminating is applied on the query trees in Fig. 10b and Fig. 10c to combined the partial results and remove the redundant unmatched windows. Its runtime for the TP left-outer join is much higher than the sum of the runtimes of the windows as presented in Fig. 11 and Fig. 12. The reason for that is that when the union of the query trees in Fig. 10b and 10c is performed, the  $\theta$  condition of the TP join is ignored for the right subtree of Fig. 12. The optimizer opts for a nested


**Fig. 12: Negating Windows**

loop for its computation and this takes a huge toll on TA's runtime making NJ two orders of magnitude faster.

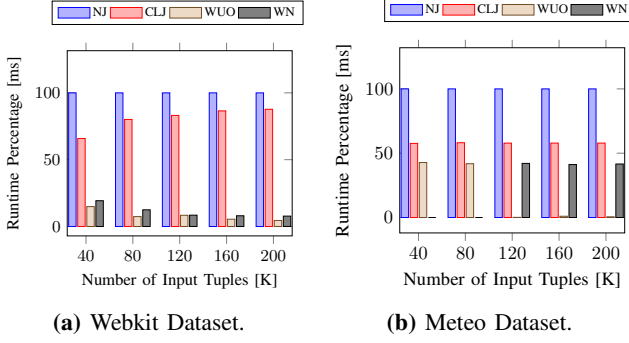

**Fig. 13: TP Left Outer-Join**

Meteo dataset contains a number of distinct values much smaller than its size, an analogy maintained in the subsets due to the use of the uniform distribution in their creation. As a result, the condition is not very selective and the runtime of both NJ and TA is higher than it was in the case of the webkit dataset. In all cases, the runtime of NJ outperforms TA by four to ten times.

#### D. Runtime Breakdown and Scalability

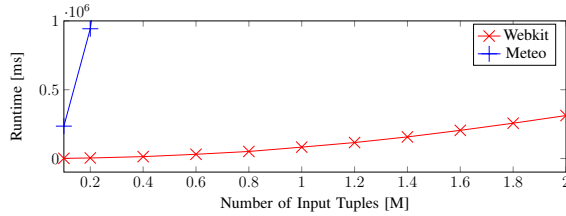
The query tree of the NJ approach (cf. Fig. 10a) consists of the nodes  $\bowtie_{\theta \wedge \theta_o}$ ,  $\mathcal{W}_{uo}$  and  $\mathcal{W}_n$  nodes. The way that the node  $\bowtie_{\theta \wedge \theta_o}$  is computed is completely determined by PostgreSQL's optimizer, given the condition applied on the non-temporal attributes. The most demanding part of the node  $\mathcal{W}_n$  is handling the tuples valid over the interval of the window. In Fig. 14, we breakdown the runtime of a TP left outer join on the percentage occupied by each node of the query tree for Webkit and Meteo dataset, respectively. As shown in the graphs, the conventional left-outer join (CLJ) occupies most of the runtime of the TP left outer join (NJ) which is more than 50% for Webkit dataset. The calls to  $LAW_U$  and  $LAW_N$ , for the computation of the nodes  $\mathcal{W}_{uo}$  and  $\mathcal{W}_n$  respectively, correspond to a small percentage of the runtime in Webkit dataset. However, they tend to be more time-consuming for Meteo dataset. This behaviour lies in the dataset characteristics and in the query performed. In meteo, the  $\theta$  condition used requests for the tuples combined to have the same metric but to refer to different stations. Measurements over all stations take place at similar times and, for multiple output intervals,

all valid tuples might contribute in the output, making the computations much more demanding.



**Fig. 14: Runtime Breakdown.** CLJ is  $\bowtie_{\theta \wedge \theta_o}$  and NJ is  $\bowtie_{\theta}^{TP}$ .

NJ is the only scalable approach integrated in PostgreSQL that can be used for the computation of all TP joins including negation. In Fig. 15, we depict the performance of NJ for the computation of a TP left outer join for larger subsets of the webkit and meteo datasets. TA is not taken into consideration, since its runtimes were already one to four orders of magnitude higher than NJ's when applied on the smaller datasets. The dataset sizes vary from 100K to 1M tuples. NJ's implementation is based on a conventional left outer join and its performance is influenced by the condition on the non-temporal attributes, since the optimizer opts for a different type of join. The selectivity of the condition applied in the webkit dataset is higher, allowing for the computation of the left outer join using a merge join. On the contrary, in the case of meteo dataset, a nested loop has to be computed. As a result, NJ scales more efficiently when applied on the webkit dataset, with its runtime being two minutes on average and always less than five minutes for datasets less than 2M.



**Fig. 15: Scalability**

## VIII. CONCLUSIONS

In this work, we proposed an approach for the computation of temporal-probabilistic joins with negation, operations that cannot currently be performed by any existing TP approach. We introduced the generalized lineage-aware temporal windows, to bind lineages and intervals and comply with the requirements of TP joins. We grouped these windows into three sets and, using these sets, we expressed the result of each TP join with negation. We implemented algorithms for the pipelined computation of all sets of generalized lineage-aware temporal windows and we integrated our approach in

the kernel of PostgreSQL. A thorough experimental evaluation reveals that our implementation is seamlessly integrated into the DBMS and outperforms existing approaches.

## REFERENCES

- [1] M. H. Böhlen, R. Busatto, and C. S. Jensen, "Point-versus interval-based temporal data models," in *Proceedings of the Fourteenth International Conference on Data Engineering, Orlando, Florida, USA, February 23-27, 1998*, 1998, pp. 192–200.
- [2] M. H. Böhlen and C. Jensen, "Sequenced Semantics," in *Encyclopedia of Database Systems*. Springer Berlin, Heidelberg, Germany, 2009, pp. 2619–2621.
- [3] D. Suciu, "Probabilistic Databases," in *Encyclopedia of Database Systems*. Springer Berlin, Heidelberg, Germany, 2009, pp. 2150–2155.
- [4] T.-Y. Wang, C. Re, and D. Suciu, "Implementing not exists predicates over a probabilistic database," in *QDB/MUD*, 2008, pp. 73–86.
- [5] K. Papaioannou, M. Theobald, and M. Böhlen, "Supporting set operations in temporal-probabilistic databases," in *ICDE*, 2018, pp. 1180–1191.
- [6] M. Dylla, I. Miliaraki, and M. Theobald, "A temporal-probabilistic database model for information extraction," *PVLDB*, vol. 6, no. 14, pp. 1810–1821, 2013.
- [7] M. Al-Kateb, A. Ghazal, A. Crolotte, R. Bhashyam, J. Chiman-chode, and S. P. Pakala, "Temporal query processing in teradata," in *EDBT/ICDT*, 2013, pp. 573–578.
- [8] N. A. Lorentzos and Y. G. Mitsopoulos, "SQL extension for interval data," *TKDE*, vol. 9, no. 3, pp. 480–499, 1997.
- [9] J. R. R. Viqueira and N. A. Lorentzos, "SQL Extension for Spatio-temporal Data," *VLDB-J*, vol. 16, no. 2, pp. 179–200, 2007.
- [10] A. Dignös, M. H. Böhlen, and J. Gamper, "Temporal alignment," in *SIGMOD*, 2012, pp. 433–444.
- [11] A. Dignös, M. H. Böhlen, J. Gamper, and C. S. Jensen, "Extending the Kernel of a Relational DBMS with Comprehensive Support for Sequenced Temporal Queries," *TODS*, vol. 41, no. 4, pp. 26:1–26:46, 2016.
- [12] D. Piatov, S. Helmer, and A. Dignös, "An interval join optimized for modern hardware," in *ICDE*, 2016, pp. 1098–1109.
- [13] L. Arge, O. Procopiuc, S. Ramaswamy, T. Suel, and J. S. Vitter, "Scalable sweeping-based spatial join," in *VLDB*, 1998, pp. 570–581.
- [14] A. D. Sarma, M. Theobald, and J. Widom, "Exploiting lineage for confidence computation in uncertain and probabilistic databases," in *ICDE*, 2008, pp. 1023–1032.
- [15] O. Benjelloun, A. D. Sarma, A. Y. Halevy, M. Theobald, and J. Widom, "Databases with uncertainty and lineage," *VLDB J*, vol. 17, pp. 243–264, 2008.
- [16] R. Fink, D. Olteanu, and S. Rath, "Providing support for full relational algebra in probabilistic databases," in *ICDE*, 2011, pp. 315–326.
- [17] R. Fink and D. Olteanu, "Dichotomies for queries with negation in probabilistic databases," *ACM Trans. Database Syst.*, vol. 41, pp. 4:1–4:47, 2016.
- [18] D. Olteanu, J. Huang, and C. Koch, "Sprout: Lazy vs. eager query plans for tuple-independent probabilistic databases," in *ICDE*, 2009, pp. 640–651.
- [19] N. N. Dalvi and D. Suciu, "Efficient query evaluation on probabilistic databases," *VLDB J*, vol. 16, no. 4, pp. 523–544, 2007.
- [20] —, "The dichotomy of probabilistic inference for unions of conjunctive queries," *J. ACM*, vol. 59, no. 6, pp. 30:1–30:87, 2012.
- [21] D. Olteanu and J. Huang, "Using OBDDs for efficient query evaluation on probabilistic databases," in *SUM*, 2008, pp. 326–340.
- [22] R. Fink, J. Huang, and D. Olteanu, "Anytime approximation in probabilistic databases," *VLDB J*, vol. 22, no. 6, pp. 823–848, 2013.
- [23] R. Fink and D. Olteanu, "On the optimal approximation of queries using tractable propositional languages," in *ICDT*, 2011, pp. 174–185.
- [24] W. Gatterbauer and D. Suciu, "Oblivious bounds on the probability of boolean functions," *TODS*, vol. 39, no. 1, p. 5, 2014.
- [25] —, "Approximate lifted inference with probabilistic databases," *PVLDB*, vol. 8, no. 5, pp. 629–640, 2015.
- [26] D. Olteanu, J. Huang, and C. Koch, "Approximate confidence computation in probabilistic databases," in *ICDE*, 2010, pp. 145–156.
- [27] A. Dignös, M. H. Böhlen, and J. Gamper, "Overlap interval partition join," in *SIGMOD*, 2014, pp. 1459–1470.
- [28] F. Cafagna and M. H. Böhlen, "Disjoint interval partitioning," *VLDB J*, vol. 26, no. 3, pp. 447–466, 2017.